



## **Proeftuinplan: Meten is weten!**

Toetsen: hoog, laag, vooraf, achteraf? Werkt het nu wel?

Middels een wetenschappelijk onderzoek willen we onderzoeken wat de effecten zijn van het verhogen cq. verlagen van toetspercentages of vooraf cq. achteraf toetsen op de kwaliteit van door zorgaanbieders ingediende aanvragen, blijkend uit veranderingen in het percentage conform getoetste aanvragen. De uitkomst van dit onderzoek kan gebruikt worden bij de visie en beleidsvorming op toezicht en handhaving, in het kader van de veranderende rol van het CIZ

## Inhoudsopgave

<b>1.</b>	<b>Inleiding .....</b>	<b>3</b>
<b>2.</b>	<b>Onderzoek.....</b>	<b>3</b>
	• <i>Vooraf of achteraf toetsen van Hits: .....</i>	<i>4</i>
	• <i>Toetsingspercentage hoog of laag bij Sips.....</i>	<i>4</i>
	• <i>Toetsing afhankelijk maken van record aanbieder (voor zowel Hits als Sips)...</i>	<i>4</i>
	• <i>Reguliere aanvragen.....</i>	<i>5</i>
<b>4.</b>	<b>Praktische implementatie.....</b>	<b>5</b>
	• <i>Toetsing vooraf/ achteraf: .....</i>	<i>5</i>
	• <i>Verhogen/ verlagen toetsingspercentage.....</i>	<i>6</i>
	• <i>Verdeling van de zorgaanbieders over de groepen .....</i>	<i>6</i>
	<b>Appendix A - Simulatiestudie naar opzet onderzoek effect toetsingspraktijk.....</b>	<b>8</b>

# Proeftuin Meten is Weten; Onderzoek Ciz-VU

## 1. Inleiding

Het CIZ moet jaarlijks ongeveer een miljoen aanvragen verwerken, dit gebeurt in 10 regionale kantoren. Deze aanvragen hebben betrekking op de vanuit de AWBZ gefinancierde zorg. Dit betreft vaak langdurige zorg voor persoonlijke verzorging, verpleging, begeleiding, behandeling en verblijf.

Van alle aanvragen die het CIZ ontvangt is zo'n 5% een Herindicatie via Taakmandaat (hits), 34% een Standaard Indicatieprotocol (sips) en 7% indicatiemeldingen. Dit zijn alle aanvragen die aan de "voorkant" direct afgehandeld worden. In aantallen is dit jaarlijks respectievelijk zo'n 45.000 hits, 340.000 sips en zo'n 65.000 meldingen. Hits komen dus minder vaak voor dan sips, maar betreft vaak de duurdere zorgverlening. Sips beschrijven over het algemeen de goedkopere zorg.

Een deel van de aanvragen sips en hits worden steekproefsgewijs getoetst door de afdeling Externe Toetsing. Van de sips wordt zo'n 5 à 10% achteraf getoetst en van de hits wordt ongeveer 25% vooraf getoetst. De indicatiemeldingen worden in het geheel niet getoetst, maar door de zorgaanbieder zelf bepaald (dit betreft opname in een instelling van cliënten ouder dan 80 jaar). Vrijwel alle overige indicaties worden door de indicatiestellers in de units behandeld. Dit betreffen er jaarlijks nog eens zo'n 500.000.

De toetsing van de hits gebeurt gemiddeld binnen één week (norm is binnen 2 weken). Zo'n 13% van de getoetste hits is niet conform, waarna een bijstelling (naar boven of beneden) plaatsvindt. Voor zorgaanbieders met een hoog percentage conform is het toetsingspercentage recentelijk verlaagd van 25% tot minimaal 21%. De ambitie is om 95% van de aanvragen conform toetsing te krijgen.

Sips worden achteraf getoetst, maar er is geen sanctie als een sip niet conform geacht wordt. Er vindt dan hooguit terugkoppeling door relatiebeheer met de zorgaanbieder plaats.

Reguliere aanvragen vormen zo'n 50% van de totale aanvragen. Deze worden onderzocht door en hierop vindt een besluit plaats door de indicatiesteller. Bij veel van deze aanvragen wordt functie en omvang zeer specifiek aangevraagd. De besluitvorming door de indicatiesteller kun je zien als een 100% toetsing vooraf.

## 2. Onderzoek

Bovenstaande praktijk laat zien dat er een aantal verschillen bestaan in toetsingspraktijk. Ten eerste is er variatie in het moment van de toetsing, vooraf of achteraf, en ten tweede varieert de hoogte van het toetsingspercentage. Dat verschilt niet alleen tussen hits, sips en reguliere aanvragen, maar ook tussen zorgaanbieders. Dit geeft aan waar de interessante beleidsvariatie ligt om te onderzoeken. Inzicht in de verschillende effecten van diverse vormen van toetsingspraktijk kan aangeven hoe de kleine geldstroom kan worden benut om de grote geldstroom (minder niet-conform getoetste aanvragen, bv. omdat de aangevraagde zorgzwaarte minder vaak te hoog is) te beïnvloeden.

Hierna volgen een aantal voorstellen om te onderzoeken. Deze hebben als doel veranderingen in het gedrag van de aanvrager te meten met betrekking tot het juist of niet juist indienen van een aanvraag. Daarmee zou je ook iets kunnen zeggen over de invloed die wij als CIZ kunnen uitoefenen op de grote geldstroom. Deze voorstellen zijn tot stand gekomen in overleg met drie onderzoekers van de VU, Birgit van Veldhuizen (Relatiebeheer en Externe Toetsing), John Braspeninx en Monique Voogt (Unit 1) en zijn voorgelegd aan Arne van Huis van Onderzoek.

- *Vooraf of achteraf toetsen van Hits:*

Voor zorgaanvragen die achteraf getoetst worden is aanpassing van bijvoorbeeld de zorgzwaarte niet mogelijk. Daarentegen kunnen aanvragen die vooraf worden getoetst wel worden bijgesteld. De doorlooptijd van een toetsing vooraf is kort zodat zorgaanbieders hiervan weinig hinder lijken te ondervinden. Een interessante onderzoeksvraag is in welke mate zorgaanbieders reageren op een wijziging in het moment van toetsing. Dit wordt in dit onderzoek onderzocht bij de Hits. Dit zou betekenen dat voor een willekeurige steekproef van zorgaanbieders (bepaalde) Hits achteraf getoetst worden, waarbij de indicatie niet bijgesteld wordt, wanneer deze niet conform is. Dit zou zich kunnen richten op een bepaalde soort aanvraag binnen de Hits. Wijzigingen in het moment van toetsing dienen naar de zorgaanbieders te worden gecommuniceerd.

- *Toetsingspercentage hoog of laag bij Sips*

Naast het moment van toetsing kan ook met het toetsingspercentage worden gevarieerd. In de afgelopen tijd zijn alleen kleine wijzigingen (naar beneden) in het toetsingspercentage van de Hits aangebracht. Een interessante onderzoeksvraag is in welke mate zorgaanbieders reageren op aanpassingen, zowel omhoog als omlaag, in de hoogte van toetsingspercentages. Dit wordt in dit onderzoek bekeken bij de Sips. Hiertoe worden zorgaanbieders ingedeeld in een groep waar weinig wordt getoetst, een groep waar normaal wordt getoetst, en een groep waar veel wordt getoetst. Bij een vierde groep kun je op prestatiebasis, waarbij prestatie gedefinieerd is in termen van percentages conform getoetste aanvragen, variëren in de hoogte van de toetsing, dit is uitgeschreven onder punt 3.

- *Toetsing afhankelijk maken van record aanbieder (voor zowel Hits als Sips)*

Onderzoeksvraag 2) zegt iets over hoe aanbieders reageren op de mate waarin getoetst wordt. Het variëren van het toetsingspercentage kan echter ook een andere prikkel hebben. Als het toetsingspercentage van zorgaanbieders die het vaak goed doen lager wordt, dan hebben zorgaanbieders een prikkel om zorgvuldig te zijn met hun aanvragen. Een grotere fractie conform getoetste aanvragen zal dan namelijk gepaard gaan met minder bureaucratie. Een systeem van straffen en belonen kan daarmee andere effecten hebben dan het veranderen van toetsingspercentages voor alle zorgaanbieders, ongeacht prestatie. In het verleden zijn de toetsingspercentages van de hits voor een aantal zorgaanbieders op basis van prestatie naar beneden aangepast. De aanpassing in het percentage getoetste aanvragen is echter klein geweest (maximaal 4 procent punt omlaag). Bovendien zijn toetsingspercentages niet verhoogd in geval van een bepaalde fractie niet-conform getoetste aanvragen. Interessant is om te kijken naar de effecten van (grotere) aanpassingen in toetsingspercentages op basis van prestatie. In dat opzicht kan

bijvoorbeeld het toetsingspercentage met 5 procent punt worden verlaagd wanneer 95% van de aanvragen van een zorgaanbieder conform is getoetst, terwijl meer dan 15% niet-conform getoetste aanvragen een directe verhoging van het toetsingspercentage met bijvoorbeeld 10 procent punt tot gevolg heeft. Bij het naar beneden bijstellen dient wel een ondergrens van bijvoorbeeld 5% getoetste aanvragen te worden gehanteerd.

Eveneens kunnen op prestatiebasis aanpassingen in het moment van toetsing worden doorgevoerd. In het geval van hits kunnen zorgaanbieders worden beloond voor een groot percentage conform getoetste aanvragen door het moment van toetsing te veranderen naar achteraf. Voor sips kunnen de effecten worden onderzocht van een vorm van sanctiebeleid waarin een te groot percentage niet-conform getoetste aanvragen zich vertaalt naar toetsing vooraf.

- *Reguliere aanvragen*

Eventueel voor de toekomst: deze onderzoeksvraag is niet alleen interessant voor hits en sips, maar ook voor de reguliere aanvragen. Zorgaanbieders weten op dit moment dat alle reguliere aanvragen vooraf worden getoetst. Aansluitend op onderzoeksvraag 3) kan het mandaatgebied van een zorgaanbieder als beloning worden uitgebreid naar de reguliere aanvragen. Een zorgaanbieder die hoog scoort op het doen van conforme aanvragen, krijgt dan het mandaat om ook een deel van de reguliere aanvragen zelf te doen.

#### **4. Praktische implementatie**

Er is gekozen voor de volgende opzet:

- *Toetsing vooraf/ achteraf:*

Bij de Hits zullen we de zorgaanbieders indelen in drie groepen. Ten eerste een controle groep waarin volgens de normale procedure vooraf getoetst wordt, dus een steekproef van 25% van de aanvragen. Ten tweede, een groep zorgaanbieders waarin een steekproef van 25% achteraf getoetst wordt (zonder aanpassing van het besluit). En ten derde, een groep zorgaanbieders waarvan achteraf getoetst wordt (zonder aanpassing van het besluit) als het conform-percentage boven de 92% blijft en anders wordt vooraf getoetst.

Door uitkomsten in de eerste en tweede groep met elkaar te vergelijken, kunnen we het hierboven genoemde onderzoek 1) uitvoeren. En door de eerste en de derde groep met elkaar te vergelijken, kunnen we voor de Hits onderzoek 3) uitvoeren. De benodigde data voor de kwantitatieve analyse zijn dat we voor elke aanvraag moeten weten, (a) op welke datum deze aanvraag gedaan is, (b) of er getoetst is, (c) of de aanvraag conform of niet-conform beoordeeld is, (d) in het geval van niet-conform op welke criteria de aanvraag niet conform was, (e) welke bijstelling er plaatsgevonden heeft. Voor het onderzoek is het ook handig om gegevens te hebben van de periode voorafgaande aan het experiment.

Het meest gecompliceerd is het bijstellen van de beslissing om vooraf of achteraf te toetsen bij zorgaanbieders die ingedeeld worden in de derde groep. De suggestie is om tweewekelijks de beslissing om vooraf of achteraf te toetsen bij te stellen. Het conform-percentage wordt berekend over de afgelopen periode van twee maanden. Een belangrijk aandachtspunt is dat zorgaanbieders in groep 2 (achteraf toetsen) en groep 3 (toetsen

afhankelijk van prestaties) vooraf geïnformeerd worden. De zorgaanbieders weten zo waar zij aan toe zijn in deze pilot.

- *Verhogen/ verlagen toetsingspercentage*

In het onderzoek met de Sips zullen we kijken naar alle Sips, in de categorieën laag risico en hoog risico. Bij de Sips zullen we de zorgaanbieders onderverdelen in 4 groepen, waarbij moet worden opgemerkt dat toetsing altijd achteraf plaatsvindt. Ten eerste een groep waarbij 6% (laag risico) en 16% (hoog risico) van de aanvragen getoetst wordt, dit is de controle groep. De tweede groep bestaat uit zorgaanbieders die minder streng getoetst zullen worden, het toetsingspercentage wordt voor deze groep verlaagd naar 2% en 10%. Bij de derde groep wordt het toetsingspercentage juist verhoogd naar 10% en 26%. Deze percentage zijn gekozen zodat het gemiddelde toetsingspercentage 6% (16%) blijft. Voor zorgaanbieders die in de vierde groep ingedeeld worden, wordt het toetsingspercentage aangepast aan hun prestaties. Zorgaanbieders die in deze laatste groep vallen zullen vooraf geïnformeerd worden over het experiment.

Door de aanvragen van zorgaanbieders in de eerste drie groepen met elkaar te vergelijken, kan een antwoord worden gegeven op onderzoeksvraag 2) en door de eerste groep te vergelijken met de vierde groep kan onderzoek 3) gedaan worden. De soort data die nodig is, is vergelijkbaar als de data die nodig zijn voor de Hits. Het aanpassen van de toetsingspercentages in de vierde groep zal twee wekelijks gebeuren. Voor elke aanvraag die in de afgelopen twee weken niet-conform is getoetst wordt het toetsingspercentage met 3%-punt verhoogd, terwijl voor elke aanvraag die in de afgelopen twee weken conform is getoetst het toetsingspercentage met 0.2%-punt wordt verlaagd. Het toetsingspercentage blijft dan ongeveer stabiel als 94% van de aanvragen conform wordt getoetst. In het bijstellen van de toetsingspercentages dient een ondergrens van 2% en een bovengrens van 26% te worden gehanteerd.

- *Verdeling van de zorgaanbieders over de groepen*

Zoals hierboven al staat beschreven verdelen we de zorgaanbieders willekeurig over vier groepen als het gaat om Sips en 3 groepen als het gaat om Hits. Aan het onderzoek doen alleen zorgaanbieders mee die regelmatig een Hit-aanvraag en/of een Sip-aanvraag doen in de eerder genoemde protocollen. We gebruiken alleen de zorgaanbieders die in totaal 4 of meer aanvragen hebben gedaan in April 2012. Dit gaat bij de Hits om 316 zorgaanbieders en bij de Sips om 254 zorgaanbieders. Om te zorgen dat alle groepen even groot zijn, worden voor de Sips 2 van de 254 zorgaanbieders (random) niet ingedeeld in een van de vier groepen. Om diezelfde reden gebruiken we voor de Hits 315 van de 316 zorgaanbieders. Bij het randomiseren in groepen hebben we rekening gehouden dat het gemiddeld aantal Hits and Sips-aanvragen tussen de groepen ongeveer gelijk moet zijn.



In Appendix A laten we de resultaten van een simulatiestudie zien om de looptijd van het experiment te bepalen. Bij een simulatiestudie wordt gekeken hoe groot de kans is dat bij een bepaalde steekproefomvang een gegeven effect van het experiment gevonden wordt. Op basis van recente gegevens over de omvang van hits en sips en het aantal zorgaanbieders, vertalen we de steekproefomvang naar looptijd van het onderzoek. Voor de sips die nu in het laag risico segment vallen zal het moeilijk zijn om met een korte looptijd een klein effect te vinden. Voor de sips die in het hoog risico segment vallen is een looptijd van 5 maanden waarschijnlijk voldoende. En dit is ook genoeg voor de hits. Het voorstel is daarom om het experiment 5 maanden te laten lopen met mogelijk een uitloop van een extra maand. Deze extra maand kan ook gebruikt worden als blijkt dat het experiment een korte aanloop in het begin nodig heeft. In dat geval zou kunnen worden besloten de eerste maand van het onderzoek niet in de analyses te betrekken.

### ***Appendix A - Simulatiestudie naar opzet onderzoek effect toetsingspraktijk***

Het belang van een simulatiestudie is om te bepalen hoeveel data nodig is om statistisch een betrouwbaar antwoord te geven. Bij veel empirisch onderzoek wordt ernaar gestreefd om een power van minimaal 0.6 a 0.7 te krijgen. De onderstaande tabellen geven een ondergrens aan de power, omdat de gebruikte statistische modellen conservatief zijn en minder gegevens gebruiken dan beschikbaar. De tabellen kunnen daarom gebruikt worden bij het bepalen van de schaal en looptijd van elk experiment.

Bij de Sips zijn momenteel twee toetsingsprotocollen, normale toetsing (6%) en striktere toetsing (16%). We gaan er vanuit dat er 240 zorgaanbieders zijn die verdeeld worden over vier groepen, dus dat betekent 60 aanbieders per groep. Voor Sips die vallen onder de normale toetsing hebben de aanbieders gemiddeld 16 aanvragen per maand. Voor Sips die vallen onder de striktere toetsing is dat 44. We gaan uit van een conformpercentage van 80%. De onderstaande tabel geeft aan hoe groot de kans is dat we een significant effect vinden als het werkelijke effect van een verhoging van het toetsingspercentage (naar 10% en 25%) gelijk is aan 2.5%, 5%, 7.5% of 10% punt. We laten dit zien voor verschillende looptijden van het onderzoek

Looptijd	2.5%	5%	7.5%	10%
<i>Normale toetsing</i>				
1 maand	0.08	0.13	0.21	0.37
2 maanden	0.07	0.14	0.35	0.56
3 maanden	0.08	0.20	0.47	0.71
4 maanden	0.10	0.27	0.56	0.87
5 maanden	0.11	0.33	0.67	0.94



6 maanden	0.15	0.43	0.79	0.96
-----------	------	------	------	------

---

*Strikte toetsing*

1 maand	0.14	0.53	0.87	0.99
2 maanden	0.28	0.84	0.99	1.00
3 maanden	0.42	0.95	1.00	1.00
4 maanden	0.55	0.98	1.00	1.00
5 maanden	0.65	1.00	1.00	1.00
6 maanden	0.73	1.00	1.00	1.00

---

Ook bij de hits zijn er 315 zorgaanbieders die voldoende groot zijn om mee te doen aan het onderzoek, maar deze worden onderverdeeld in drie groepen. Gemiddeld doet een zorgaanbieder per maand 15 aanvragen, waarvan 25% getoetst wordt. We gaan uit van een conformpercentage van 85%. Evenals bij de Sips laten we in de tabel hieronder zien wat de power is als de andere wijze van toetsing een effect heeft op het conformpercentage van 2.5%, 5%, 7.5% of 10% punt.

# aanvragen	<b>2.5%</b>	<b>5%</b>	<b>7.5%</b>	<b>10%</b>
1 maand	0.10	0.25	0.55	0.85
2 maanden	0.15	0.55	0.89	1.00
3 maanden	0.23	0.74	0.98	1.00
4 maanden	0.29	0.84	0.99	1.00
5 maanden	0.34	0.88	1.00	1.00
6 maanden	0.39	0.93	1.00	1.00

---